

REMARKS

This is in response to the Final Office Action mailed on February 15, 2008. In this Office Action, claims 1-30 were rejected. In this response, claims 2-5, 18 and 19 have been cancelled while independent claims 1 and 17 have been amended.

Amendment are Appropriate for Entry

Applicants first note that all proposed amendment made in this Amendment After Final are appropriate for entry according to 37 CFR 1.116. 37 CFR 1.116 states that an amendment after final may be made canceling claims. Furthermore, 37 CFR 1.116 states an amendment after final presenting rejected claims in better form for consideration on appeal may be admitted. Applicants submit that the present amendments are provided for clarity and do not change the scope of the claims. As such, Applicant submits that the present amendments do not require a new search and respectfully request that the amendments are entered.

In conclusion, it is respectfully pointed out that all the presently submitted amendments essentially amount to the relocation of dependent claim features into independent claims. Thus, the amendments are entirely appropriate for entry after final rejection. Consideration and entry of the amendments are respectfully solicited.

Claim Rejections

Claims 1, 6-16

Claims 1-2 and 6-16 were rejected under 35 USC 102(b) as being anticipated by Shinyama et al. (hereinafter "Shinyama"). In this response, claim 1 has been amended to include the limitations from claim 2.

Applicant's Specification describes a textual alignment system that operates on large sets of sentences by extracting mappings between words or phrases based on a holistic examination of sentences in the set. Lexical correspondences between the sentences in the sets are learned. The Specification describes an alignment algorithm that establishes links between different, but parallel pieces of information. As shown in FIG. 4, this can be illustrated through

lines that connect various words and phrases together. The nature of the alignment process is such that as more data is acquired, the links between incorrect correspondences will fade.

The Office Action indicated that claim 1 is anticipated by Shinyama. The Office Action indicated that the limitation “using statistical textual alignment to align words in the text segments in the set” is equivalent to the statistical entity tagging system of Shinyama. The Shinyama reference itself says very little about the “statistical entity tagging system.” On page 3 it says “We mark all NEs using an statistical NE tagging system.” During its discussion of the statistical NE tagging system, Shinyama references paper [7], “Named Entity Extraction Based on a Maximum Entropy Model and Transformation Rules,” which actually teaches the statistical named entity tagging system. The Abstract of this referred to reference indicates that the paper “describes named entity (NE) extraction based on a maximum entropy (M.E.) model and transformation rules.” This paper does not teach statistical textual alignment, but rather describes how named entities are extracted. Therefore, the statistical named entity tagging system which is taught in Shinyama extracts named entities based on maximum entropy and transformation rules, but does not use “statistical textual alignment to align words in the text segments in the set,” as claimed in claim 1.

The Office Action indicated the use of named entities as taught by Shinyama on pg. 3 of 6, section 3.2, step 4 is equivalent to the limitation “identifying the paraphrase relationships based on the aligned words” (emphasis added). Arguendo, assuming the Office Action is correct in its assertion that Shinyama’s statistical named entity tagging system teaches the limitation “using statistical textual alignment to align words in the text segments in the set”, there is no teaching of the limitation “identifying the paraphrase relationships based on the aligned words” (emphasis added). Page 2 of Shinyama states “Each IE pattern has slots which can be filled by NEs... NEs in these patterns are generalized into slots which hold the type of NEs... [we] apply these obtained patterns to the articles itself, and then find paraphrases only among those which matches to any of patterns. This means we find paraphrases among these IE patterns. Actually this is done by linking two IE patterns as paraphrases” (emphasis added). Therefore, Shinyama teaches two IE patterns are linked as paraphrases. In this case, the named

entities do not identify the paraphrase relationships as the Office Action asserts. Rather, the IE patterns are used to identify paraphrase relationships. Therefore, the limitation “identifying the paraphrase relationships based on the aligned words” cannot be met even if the Office Action is correct in its assertion “using statistical textual alignment to align words in the text segments in the set” is taught by Shinyama since IE patterns are used to identify the paraphrase relationships instead of the named entities. It is therefore submitted that for at least the reasons discussed above, claim 1 is in form for allowance over the cited prior art. It is also submitted that claims 6-16 are in form for allowance as well due to their dependence on allowable independent claim 1.

Claims 17, 20-29

The Office Action indicated that claim 17 was rejected under 102(b) by Shinyama in view of Gibson. In this Response, claim 17 has been amended to include limitations from claims 18 and 19. Claim 17 has been amended to:

“A paraphrase processing system, comprising
a textual alignment component configured to receive a set of text segments and identify paraphrase relationships between words in the set of text segments based on alignment of the words, wherein the textual alignment component is configured to generate an alignment model based on a heuristic alignment of the words, and wherein the textual alignment component is further configured to identify paraphrase relationships based on alignments of multi-word phrases in the set of text segments” (emphasis added).

The Office Action indicated the limitations “wherein the textual alignment component is configured to generate an alignment model based on a heuristic alignment of the words, and wherein the textual alignment component is further configured to identify paraphrase relationships based on alignments of multi-word phrases in the set of text segments” (emphasis added) were rejected under 35 U.S.C. 103(a) as unpatentable over Shinyama in view of Gibson.

The Office Action indicated Gibson's BLAST heuristic taught the limitation "an alignment model based on a heuristic alignment of words." In sharp contrast to claim 17, Gibson's BLAST heuristic scans the database for words which comprise proteins and DNA. For example, when a novel stretch of nucleic acid, e.g., DNA, or protein is first sequenced, its sequence is typically compared against a database of known DNA and protein information to provide a preliminary indication of the new DNA or protein's function. Gibson's BLAST heuristic aligns words consisting of protein and DNA information but clearly does not generate an alignment model based on a heuristic alignment of the words.

Gibson's BLAST heuristic which aligns "words" consisting of DNA and protein symbols, and Shinyama's textual paraphrase acquisition are from vastly different arts. Some confusion has ensued in applying the holding of the recent Supreme Court decision in KSR with reference to analogous arts as opposed to reason to combine references. KSR International Co. v. Teleflex Inc. (KSR), 550 U.S. _____, 82 USPQ2d 1385 (2007), (MPEP §2141). KSR addressed reason to combine references, but did not address the separate, prerequisite issue under Graham of evaluating analogous arts to determine the proper scope of relevant prior art; and KSR was decided in a context in which the cited references were all clearly from the same narrowly specified field of art as the claimed invention, i.e. the narrowly specific art of sensors associated with gas pedals for motor vehicles. Analogous arts were simply not an issue in KSR.

The references cited in this case are far from analogous arts, and KSR is not applicable in contradicting the fact that they not combine to support a reason for combining.

The proposed reasons to combine provided in the Office action also clearly exceed the scope of reasons to combine provided by the Supreme Court in KSR. The Supreme Court in KSR said a person of ordinary skill in the art (POSA) "has good reason to pursue the known options within his or her technical grasp" when "[1] there is a design need or market pressure to solve a problem and [2] there are a finite number of identified, predictable solutions" (KSR, slip op. at 17). The Court of Appeals for the Federal Circuit has since interpreted the second element of this to mean "a finite, and in the context of the art, small or easily traversed, number of options", where subject matter outside that small, easily traversed number of options does not "support an

inference of obviousness”. *Ortho-McNeil Pharma. v. Mylan Labs*, slip op. at 9-10 (Fed. Cir. 2008). In *Ortho-McNeil*, the Federal Circuit found the claimed invention non-obviousness at several levels, where a POSA would not even have been likely to start where the inventor did; the POSA also faced too many unpredictable variables without a given reason to select among them as the inventor had; and the POSA would have had to explore properties far afield from the inventor’s initial purpose.

The recitation in the Office action of illustrative advantages that would be provided by the subject matter of the present claims, does not indicate any design need or market pressure to have brought those advantages to a paraphrase processing system; it does not show that the various elements of the disparate cited references were drawn from only a finite, and in the context of the art, small or easily traversed number of options; it does not show how a person of ordinary skill in the art would have selected those elements from among the many variables available throughout all possible arts that are at most as tenuously analogous to the textual paraphrase acquisition of the cited primary reference; and it does not indicate how a person of ordinary skill and creativity in the art would have conceived of combining elements from all the disparate cited references without exploring properties far afield from the inventor’s initial purpose.

Therefore, the proposed reasons for combining the cited references far exceed the proper scope of reasons to combine under KSR, and do not support a valid showing that the present claims are obvious.

Moreover, even when Shinyama and Gibson are combined, they do not render obvious each claimed limitation. For example, the paraphrase system in Shinyama relies on IE patterns to generate paraphrases as discussed above. The heuristic in Gibson aligns words consisting of protein and DNA information. Even in combination, these references do not render obvious a recitation “wherein the textual alignment component is configured to generate an alignment model based on a heuristic alignment of the words, and wherein the textual alignment component is further configured to identify paraphrase relationships based on alignments of multi-word phrases in the set of text segments” as claimed. Therefore, independent claim 17 and

its dependent claims 20-29 are not rendered obvious by the combination of Shinyama and Gibson.

Claim 30

The Office Action indicated that claim 30 was rejected in view of Shinyama and under the same rational as claims 17 and 29.

The Office Action indicated that claim 30 was rejected as being unpatentable over Shinyama. The Office Action cited to the TF/IDF based method on pg. 2 of Shinyama to teach the limitation “a textual alignment component configured to receive a plurality of text segments and identify paraphrase relationships between words in the text segments based on alignment of the words.” The TF/IDF based method in Shinyama is used to find pairs of articles which report the same event. This is markedly different than using textual alignment to identify paraphrase relationships between the text segments in the set. The TF/IDF based method discussed in Shinyama, is used for categorization of newspaper articles, not for a method of textual alignment to identify paraphrase.

Furthermore, as discussed above, Page 2 of Shinyama states “Each IE pattern has slots which can be filled by NEs... NEs in these patterns are generalized into slots which hold the type of NEs... We apply these obtained patterns to the articles itself, and then find paraphrases only among those which matches to any of patterns. This means we find paraphrases among these IE patterns. Actually this is done by linking two IE patterns as paraphrases” (emphasis added). Therefore, Shinyama teaches two IE patterns are linked as paraphrases. In this case, the TF/IDF method does not identify the paraphrase relationships as the Office Action asserts. Rather, the IE patterns are used to identify paraphrase relationships. Therefore, the limitation “a textual alignment component configured to receive a plurality of text segments and identify paraphrase relationships between words in the text segments based on alignment of the words” (emphasis added) is not met by Shinyama.

The Office Action indicated the limitation “a paraphrase generator, receiving a textual input and generating a paraphrase of the textual input based on the paraphrase

relationships” was taught by Shinyama at pg. 4 of 6, section 6. Shinyama indicates that pairs of paraphrases can be built. There is no teaching that these paraphrase relationships are used in a paraphrase generator by receiving a textual input and generating a paraphrase of the textual input based on the paraphrase relationships. The particular portion of Shinyama the Office Action cites to on page 4, section 6 states “we ran the paraphrase acquisition system for each pair or articles, and finally got total 136 pairs of paraphrases (a link between two IE patterns).” Although this teaches that paraphrases are built, it does not teach the limitation “a paraphrase generator, receiving a textual input and generating a paraphrase of the textual input based on the paraphrase relationships” (emphasis added). Reconsideration and allowance of claim 30 is respectfully requested.

Conclusion

In summary, it is respectfully submitted that claims 1, 6-17 and 20-30 are in allowable form. Consideration of these claims as presently amended and favorable action is respectfully solicited.

The Director is authorized to charge any fee deficiency required by this paper or credit any overpayment to Deposit Account No. 23-1123.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: /Christopher L. Holt/
Christopher L. Holt, Reg. No. 45,844
900 Second Avenue South, Suite 1400
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312

CLH:SD:rkp